# Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms

# Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms

**Roberta Bergero[1], Jim Gardner[1], Beth Bader[1], Lengxob Yong[2], Deborah Charlesworth[1]**

[1]University of Edinburgh, [2]University of Exeter

Submitted to Proceedings of the National Academy of Sciences of the United States of America

**It is often stated that polymorphisms for mutations affecting fitness of males and females in opposite directions (sexually antagonistic, or SA, polymorphisms) are the main selective force for the evolution of recombination suppression between sex chromosomes. However, empirical evidence to discriminate between different hypotheses is difficult to obtain. We report genetic mapping results in lab-raised families of the guppy (*Poecilia reticulata*), a sexually dimorphic fish with SA polymorphisms for male coloration genes, mostly on the sex chromosomes. Comparison of the genetic and physical maps shows that crossovers are distributed very differently in the two sexes (heterochiasmy), and in male meiosis are restricted to the termini of all four chromosomes studied, including chromosome 12, which carries the sex-determining locus. Genome resequencing of male and female guppies from a population also indicates sex linkage of variants across almost the entire chromosome 12. More than 90% of the chromosome carrying the male determining locus is therefore transmitted largely through the male lineage.A lack of heterochiasmy in a related fish species suggests that it originated recently in the lineage leading to the guppy. Our findings do not support the hypothesis that suppressed recombination evolved in response to the presence of SA polymorphisms. Instead, a low frequency of recombination on a chromosome that carries a male-determining locus and has not undergone genetic degeneration has probably facilitated the establishment of male-beneficialcoloration polymorphisms.**

crossing over | sexual antagonism | sex chromosomes | guppies | genetic maps



## Introduction

In a diversity of organisms, the sex chromosome pair has evolved suppressed recombination, sometimes extending across the entire chromosome (1-3, as recently reviewed in 4, 5). One prominent theory for the repeated, independent evolution of suppressed recombination is that sexually antagonistic (SA) factors at loci closely linked to a sex-determining locus establish polymorphisms, with one allele benefitting males and becoming associated with the male-determining allele at the sex-determining locus, while another allele that is favoured in females is associated with the alternative allele at the sex-determining locus (6). Such linkage disequilibrium (LD) between the alleles at the SA locus and the sex-determining region generates selection for closer linkage between the two loci (3). Although this model is plausible, and could be important in the evolution of non-recombining regions of sex chromosomes (7), the SA polymorphism hypothesis is difficult to test, because, if recombination between the SA gene and the sex-determining region has become suppressed, the entire region co-segregates in genetic crosses. This makes it difficult to detect the existence of separate genes with SA polymorphisms within the region, so evidence supporting this evolutionary process is scarce. An important source of information comes from the observation of so-called "evolutionary strata" in which the sex chromosome pair stopped recombining at different times during their evolution. This is inferred from differences in sequence divergence detected in vertebrates, including mammals (8, 9), birds (10), and the threespine stickleback (11, 12), and in some plants (13). Such strata are consistent with the SA polymorphism hypothesis for recombination suppression, but are not definitive evidence, because other evolutionary situations can potentially select for reduced recombination around sex-determining loci (5, 14, 15).

The fish *Poecilia reticulata* (the Trinidadian guppy) is particularly suitable for studying whether SA polymorphisms indeed lead to recombination suppression, because SA selection has been shown to be ongoing in natural guppy populations (16). Male coloration factors are polymorphic within populations of guppies in many rivers in the Northern Range of mountains in Trinidad (17-20). Conspicuous male coloration is favoured by sexual selection, but also increases visibility to predators. Male coloration traits are therefore sexually antagonistic: in males, the risk of predation may be outweighed by the mating advantage (particularly in upriver populations, where predation is less severe due to waterfalls preventing the main predatory fish moving from down-river sites), whereas such traits give females no advantage (21).

Genetic studies (reviewed by 15,16) show that 79% of factors controlling guppy male coloration are concentrated on the chromosome pair that carries the sex-determining locus. This chromosome, however, represents only around 4% of the physical genome (22-24). Of the sex-linked male coloration factors, slightly more than half are fully sex-linked, while the others are partially sex-linked, being located at most about 10 centiMorgans (cM) from the sex-determining locus, and these are expressed only in males (16). Although SA selection clearly occurs in natural guppy populations, negative frequency dependent selection also contributes to maintaining these SA polymorphisms, as rare male morphs have advantages in mating (25) and higher survival (21).

### Significance

**Sexual dimorphism is common in animals, and often involve mutations that improve one sex, but harm the opposite sex. However, very little empirical evidence exists about such conflicts and their evolution, including sex chromosome evolution. We report genetic and genomic analyses in a fish, the guppy, whose males have variable, bright coloration patterns that are beneficial during courtship, but increase predation, therefore harming females. We found that genetic recombination events are strongly restricted to chromosome tips in males, but not females, and that this recombination difference between the sexes may have evolved recently in the guppy lineage. This sex difference ensures that male beneficial mutations are rarely transmitted to females, explaining their observed enrichment on the guppy's male-determining chromosome.**

**Reserved for Publication Footnotes**

**Table 1. . Summary of genetic mapping families. The most distal markers mapped in male meiosis showing complete sex linkage are indicated for each family (i.e. with map position 0 in the male maps in Fig. 1). These markers are GT350 and GT369 with genome positions of 24.5 Mb and 25.3 Mb, respectively. The slightly more distal marker GT369 was not informative and could not be mapped in families LAH and QLPB1.**

| Family name | River | Predation level in source population | Progeny number | Most distal fully sex-linked marker in males |
|---|---|---|---|---|
| LAH | Aripo | High | 42 | GT350 |
| GHP3 | Guanapo | High | 45 | GT369 |
| QHPB4 | Quare | High | 35 | GT369 |
| QHPG5 | Quare | High | 23 | GT369* |
| QHPG3 | Quare | High | 33 | GT369 |
| ALPB1 | Aripo | Low | 52 | GT369 |
| ALPB2 | Aripo | Low | 87 | GT369 |
| QLPB1 | Quare | Low | 129 | GT350 |

*a recombinant male individual was found carrying the sire's X-linked alleles at seven centromere-proximal markers, and the sire's Y-linked alleles from marker *cyclAC* to the terminus. The crossover event was therefore proximal to 21.3 Mb.



The enrichment of SA coloration factors on the sex chromosome pair suggests that a large region behaves as Y linked, and that SA polymorphisms that arose on this chromosome may subsequently have evolved either male-specific expression, or recombination suppression leading to complete sex linkage for some of the factors, the two expected routes by which conflicts between the sexes can be resolved (26).

Two kinds of studies in guppies have inferred that the association between coloration factors and the sex-determining locus differs between populations. First, treating females with testosterone to induce male development reveals that females from populations in low-predation locations carry male coloration factors more often than do females from high-predation locations (17, 27). Moreover, the results of experiments moving fish from high-predation locations to predator-free ones suggested that this association can change rapidly in the absence of predators (18). Testosterone treatment produces changes only in females carrying coloration factors. If coloration factors are at lower frequencies in populations from high- than low-predation sites, fewer females from the former are expected to exhibit changes, compared with the latter. To our knowledge, the only direct, genetic data supporting a difference in linkage are for the partially sex-linked Sb factor (17). Second, a recent genome sequencing study also suggested a recent change in recombination, but in the opposite direction; up to half of the guppy sex chromosome pair was inferred to be fully sex-linked in upstream populations, whereas in downstream, high-predation populations only an old non-recombining stratum was found (28).

In order to test for the SA polymorphism hypothesis and to shed light on contradictions emerging from previous research on guppies, we studied Trinidadian guppies using genetic and population genomic approaches that could reveal recent change in recombination. Our genetic mapping revealed very different crossover patterns in the two sexes. In male meiosis, crossing over was detected only at the termini of all chromosomes tested (chromosome 12 and three autosomes). The sex chromosome is therefore clearly not unique in this respect, but, because it carries the male-determining locus, a large region that includes this locus will be transmitted exclusively through the male lineage, so that wholly or largely Y-specific variants can be maintained (29). For chromosome 12, our population genomic results strongly suggest that crossing over is very rare across most of the chromosome, with more than 90% of its gene content showing detectable LD with the sex-determining locus.

Since heterochiasmy is genome-wide in the guppy, it probably evolved before LG12 became the sex-determining chromosome, and may not have evolved under selection generated by SA polymorphisms. Instead, it may be similar to situations in which sex-determining genes evolved in non-recombining genome regions, including pericentromeric regions of low recombination, where no evolutionary change in recombination rate is required to generate a fully sex-linked region (30). If so, rather than causing the evolution of a fully sex-linked region, the enrichment of the guppy sex chromosomes with the SA polymorphisms present in populations today may be a consequence of the rarity of recombination throughout most of this chromosome, not its cause.

## Results

*Genetic mapping*

We estimated separate genetic maps in male and female meiosis, and compared their crossover patterns, in full-sib F1 families from individuals from within-population crosses made using pairs of individuals from several high- and low-predation sites from the Aripo, Guanapo and Quare rivers (Table 1, Table S1). We chose microsatellite and SNP markers (Table S2) from variants found in our genome sequences in genes from chromosome 12 in the guppy female assembly (22), and from other chromosomes (see below). Previous high-density genetic mapping showed that this chromosome carries the guppy sex-determining locus (31, 32). In at least some families, markers were informative for both male and female meiosis (Table S2), suggesting that male hemizygosity is infrequent. In male meiosis of our families, most chromosome 12 markers co-segregate with the sex phenotype, confirming that chromosome 12 carries the sex-determining locus in all our male parent fish, and therefore probably in all guppy populations.

Table 1 about here

Despite no overall difference in recombination rates between males and females being detected in a previous study based on AFLP markers (33), and little difference in total map lengths in our data, our segregation data showed a striking sex difference in crossover patterns (Fig. 1, Fig. S1). In male meiosis, recombination with the sex -determining locus was detected only for markers located in the terminal megabase of LG12. In contrast, recombinants were detected throughout chromosome 12 in female meiosis, and the order of the markers in the female genetic map generally agrees well with their order in the female genome assembly (22). Heterochiasmy was suggested by a previous study which detected rare recombinants in inter-population crosses (31). No markers in the terminal highly-recombining PAR region were mapped in that study. The strong localization of crossovers in male meiosis found here is also consistent with cytogenetic MLHI foci locations in bivalents from testis cells (24). In the families other than the Aripo LAH family (top left in Fig. 1), we mapped only enough markers to establish that the results are consistent with those from the LAH family. Importantly, the boundary between the recombining and non-recombining regions in male meiosis is always in a similar location (Table 1, Table S1).

Figure 1 about here

All guppy chromosomes are acrocentric (23). The region that recombines in male meiosis must therefore be the end distant from the centromere, so we refer to the two ends as the "centromere" and the "terminus" in what follows. For the sex chromosome pair, LG12, the terminal region is a highly recombining pseudo-autosomal region. Among 446 genotyped progeny, we observed only a single crossover separating the sex-determining locus from centromere proximal LG12 markers, in family QHPG5 (Table 1). This crossover occurred in the interval between markers AG177 (14.4 Mb) and cyclAC (21.3 Mb),

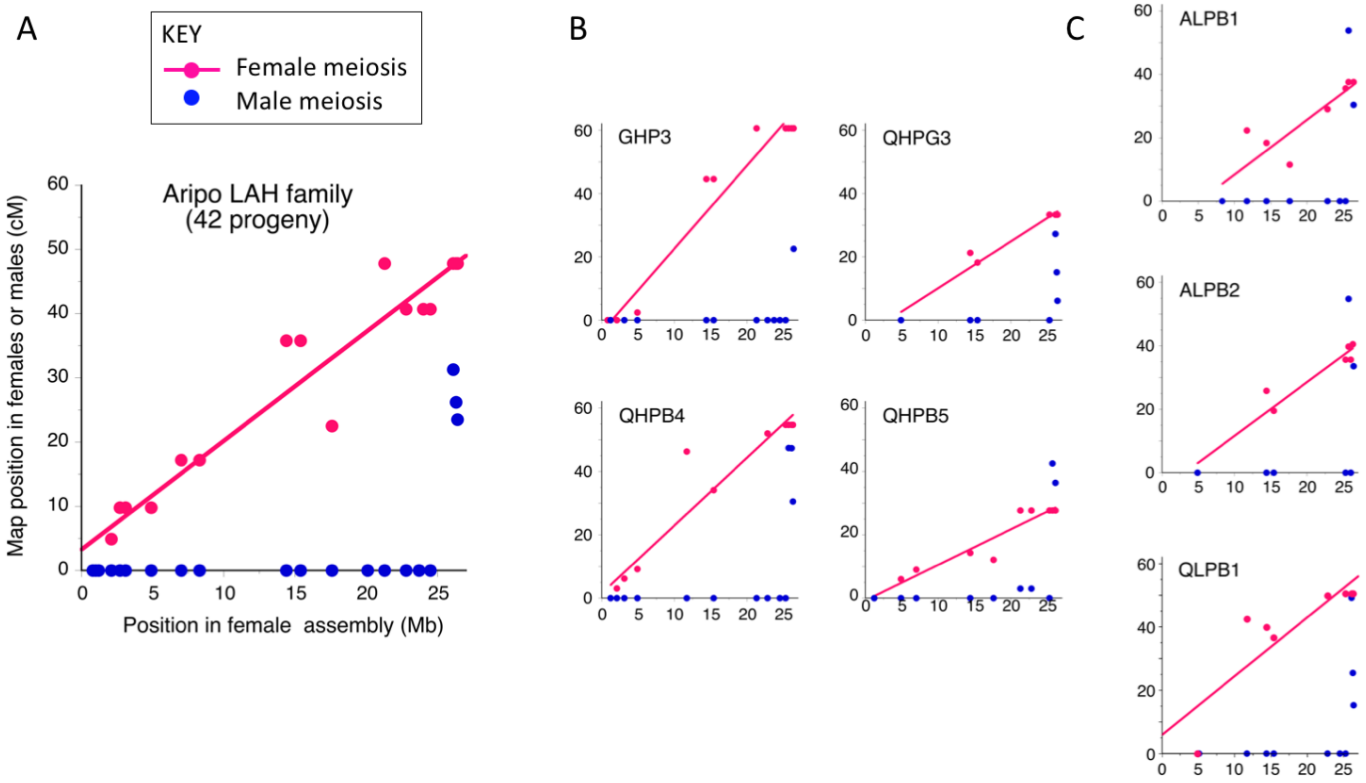


**Fig. 1.** . Genetic map distances for markers on chromosome 12 plotted against their genomic positions in the chromosome 12 female assembly available in GenBank. (A) Results from a full sib family from the Aripo captive population. Genetic mapping was also carried out for families from fish sampled directly from natural high-predation (B) and low-predation (C) populations from the Aripo, Guanapo and Quare rivers (see Methods and Table 1). In all families, the genetic map locations estimated in female meiosis increase almost linearly with the physical genomic positions, with $R^2$ values for the linear regressions close to the value of 1 for all families (Table S1). In male meiosis, the terminal region distant from the centromere, forms a PAR with a very high recombination rate.

placing the male-determining locus distal to the latter, and just proximal to the PAR.

*Population genomic tests for associations with the sex-determining locus*

Extreme localization of crossovers at the chromosome termini means that a large region of the chromosome carrying the "maleness" locus will, at most, recombine rarely with its homologous chromosome, and will be inherited largely or entirely through males, and could evolve differentiated Y haplotypes, producing evidence of Y linkage for SNPs throughout most of chromosome 12. To test this, we sequenced the complete genomes of 10 male and 6 female fish from a captive population of guppies derived from the same Aripo river population as used for the genetic mapping described above.

We analyzed $F_{ST}$ values between the sexes for individual variable sites (see Methods) to test for associations between SNPs and the sex-determining locus (34). Considerably higher $F_{ST}$ values were estimated for LG12 than for the other chromosomes (Fig. S2A), supporting its previous identification as the sex chromosome pair (23). As was previously found in a captive population of fish from a Quare river source (28, 35), SNP density was also marginally highest on LG12 (Fig. S2B).

Signature of sex linkage were indeed detected across most of chromosome 12. $F_{ST}$ analysis of 50 kb windows across chromosome 12 (see Methods) revealed clear differentiation between sequences from males and females (Fig. 2A). Consistent with this evidence for allele frequency differences between the sexes, and for male-specific variants, Tajima's D values for LG12 SNPs were higher in males than females (Fig. S3); this supports sex linkage, because males are expected to all be heterozygous at SNPs that differentiate the X and Y. However, no region had $F_{ST}$ approaching the value 0.27 expected for fully sex-linked sites with our sample size. The first three megabases of the assembly had higher male-female differentiation than elsewhere on chromosome 12 (Fig. 2A; Tajima's D values were also higher in this region, see Fig. S3). However, this region also shows linkage disequilibrium (LD, estimated as $r^2$ values, see Methods) in the pooled sample of fish of both sexes (Fig. 2C), possibly due to being close to the centromere (Fig. 2D).

Figure 2 about here

In a fully sex-linked region, high frequencies of heterozygotes specifically in males are predicted, with many sites expected to have $F_{IS}$ = -1 in males, because $F_{IS}$ measures departures from frequencies expected for autosomal genes under random mating, and negative values indicate an excess of heterozygotes. For sex-linked variants, females should have $F_{IS}$ values more positive than those in males, as their genotype frequencies should be close to the Hardy-Weinberg proportions (yielding $F_{IS}$ values near zero, or positive if the population is inbred). $F_{IS}$ analysis (see Methods) indeed supports our conclusion of sex linkage at many LG12 sites: male values are consistently lower across LG12 than those of females (and are mostly negative, indicating an excess of heterozygous SNPs, contrasting with generally positive values in females; Fig. 2B). Two regions have less $F_{IS}$ differences between the sexes. The region near the chromosome terminus corresponds to the recombining (pseudo-autosomal) region identified by our genetic mapping (see Fig. 1), and that near 11,000 Kb with negative values in both sexes could reflect duplicated sequences or an assembly error. Overall, however, the generally high $F_{ST}$ values are due to variants at intermediate frequencies that are commoner in males than females, and not to low diversity, which is a common cause of high $F_{ST}$ (36). Importantly, the finding of

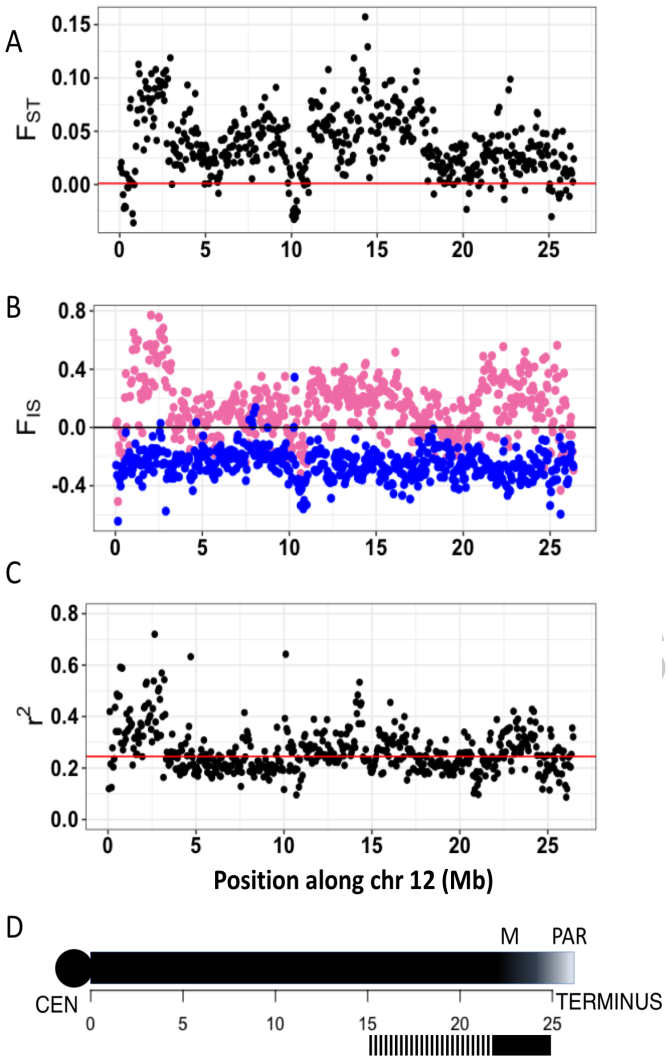


**Fig. 2.** . Evidence for associations between SNPs on chromosome 12 and the sex-determining region, based on population genetics analyses ($F_{ST}$, $F_{IS}$, LD). A total of 469,515 biallelic SNPs with genotypes for all individuals were analysed. The positions are based on the female genome assembly. **A**. $F_{ST}$ values between males and females. The values plotted are mean values in 50 kb windows; the horizontal line shows the genome-wide mean $F_{ST}$. **B**. $F_{IS}$ values for SNPs in males (blue) and females (red) separately, showing the more negative values in males in almost all 50 kb windows across LG12, except in the pseudo-autosomal region near 26 Mb. The line indicates a value of zero, the value expected under Hardy-Weinberg equilibrium. **C**. Mean linkage disequilibrium (LD) between SNPs in the same windows, estimated from diploid genotypes of both sexes using $r^2$ values. The horizontal line shows the chromosomal mean $r^2$ value. Away from the centromere end, $r^2$ decays to a value close to 0.2. As the population studied has a history of a small size at its founding (see Methods), and our sample size is small, LD is not expected to decay to very low values, even for independently segregating SNPs (74). **D**. Schematic diagram of the chromosome, showing the centromere at the left (CEN), corresponding with the region of high LD. The grey gradient of the chromosome bar indicates the very high crossing over rate at male meiosis at the tip opposite the centromere (the terminus), quickly decreasing to a very low rate across the large centromere-proximal region. The presumed location of the male determining region, based on our genotyping of one male recombinant, is indicated by "M". The positions of the old and young evolutionary strata inferred by (28), are indicated by a solid black and a striped black bar, respectively.

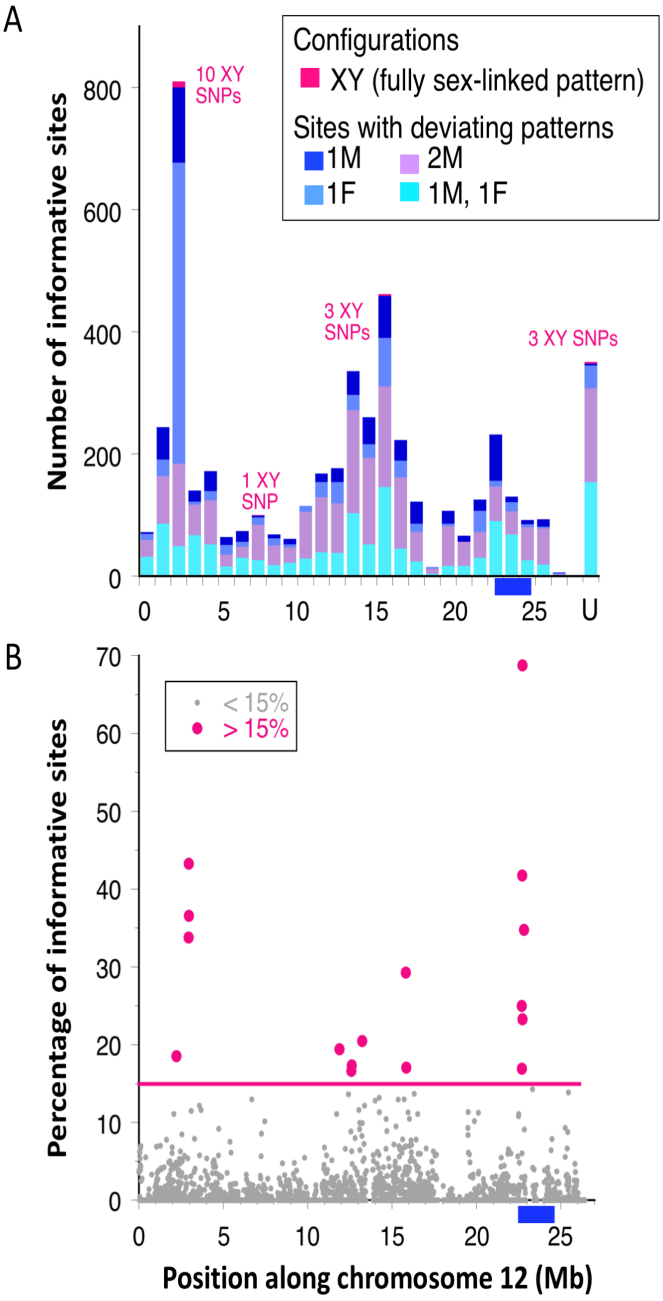


**Fig. 3.** . Evidence for restricted recombination on chromosome 12, based on sites with genotype patterns expected under sex linkage. A total of 344,848 non-singleton SNPs from chromosome 12 were analyzed in windows of 1 Mb, plus SNPs from 492 unplaced scaffolds (U) with at least 100 variable sites. Part A shows the numbers of informative sites with genotypes consistent with sex linkage. Such sites were found across most of chromosome 12. Few sites had genotype configurations compatible with full sex-linkage (all females homozygous, and all males heterozygous). Most informative sites had genotypes deviating from complete sex linkage as follows: (1M) a single male is homozygous, (2M) 2 homozygous males, (1F) one female heterozygote, and (1M, 1F) one male homozygote and one female heterozygote. Part B shows the percentage of sites in 10 kb windows with genotypes configurations that fall into any of the five categories specified above. The pink line separates the number of windows containing more than 15% (pink dots) or less than 15% (grey dots) of sex-linked sites. The blue bar under the X axis indicate the region previously suggested to be an old evolutionary stratum (28).

male-specific and male-enriched variants in regions considerably centromere-proximal to the sex-determining locus (in a distal region, see above and refs. 23 and 24) implies that much of the chromosome recombines with this locus very infrequently in males, consistent with the genetic mapping described above.

However, in total, only 14 LG12 SNPs have genotype configurations compatible with complete sex linkage (Fig. 3A). Among these, ten clustered in a 15.6 kb region between positions 2,937,371 and 2,952,959 bp in the female assembly. Three other such SNPs were in a much larger region at a considerable physical distance from this cluster (> 574 kb between positions 15,268,306 and 15,843,036), and one was located at position 7,677,458. Our sample size does not permit us to conclude firmly that any SNP is fully sex-linked, given the large number of SNPs genotyped, and the possibility of false-positives, and therefore our low estimated number of fully sex-linked SNPs on LG12 is probably an overestimate.

Figure 3 about here

We also counted the number of autosomal sites with signals of complete or partial sex linkage, to estimate the number expected by chance. Consistent with our observations for the $F_{ST}$ values (Fig. S2A), the other chromosomes show markedly smaller numbers of sites with signals of complete or partial sex linkage; in total, only 21 SNPs on the 22 chromosomes other than LG 12 had genotype configurations expected under complete sex linkage (heterozygous in all 10 males, and homozygous in all 6 females studied), and only 3 autosomes had more than two such SNPs; these are probably mostly false positives. The over-representation of LG12 is therefore striking; the random expectation, based on either the number of SNPs analyzed, or the estimated chromosome sizes, is that LG12 should carry about 4% of such SNPs, while the proportion we observed is 40% (in 10,000 trials of 35 sites assigned randomly to 23 chromosomes, the number found on a single chromosome never exceeded 7). Moreover, only three further SNPs with this pattern were detected among the unplaced scaffolds (Fig. 3A).

Detailed analysis of the genotypes of SNPs on the guppy sex chromosome pair in individual fish revealed that genotypes of most of our sample of 16 fish are frequently compatible with complete sex linkage, but a few individuals have numerous sites suggesting that recombination has occurred (Fig. S4); these individuals are discussed further below. Further evidence that recombination regularly occurs between large X- and Y-linked regions on LG12 is that variants at many sites are shared between the two sexes, and biallelic sites with male homozygotes for both alleles are abundant, which cannot occur without recombination.

However, the overall LD with the sex-determining region shows that recombination is infrequent throughout most of LG12. Genotype configurations suggesting sex linkage are particularly abundant at sites in the three regions that also have the apparently fully sex-linked SNPs (Fig. 3A), and such sites form high proportions of the SNPs in 10-kb windows in the same three regions of the assembly (Fig. 3B); these regions are therefore not produced by regions with unusually high SNP densities. However, although the region between 22 and 25 Mb (where an old stratum shared by all guppy populations has been inferred (28), and which is also identified as the sex-determining region by the recombinant described above), also includes several 10-kb windows with high proportions of sex-associated SNPs (Fig. 3B), it includes not a single SNP with the genotype configuration indicating complete sex linkage (Fig. 3A).

Overall, these results suggest that there is a very small region showing full sex-linkage, perhaps like the single gene sex-determining systems known in some other fish (37-40), and a recent new analysis in the guppy also infers a "small" ancestral fully sex-linked region (41). Coverage analysis of our sequence data, using an approach (see the Methods section) different from the previous analysis used to infer sex linkage (28), did not detect coverage differences between the sexes in this region (Fig. S5) or evidence for male hemizygosity elsewhere on LG12 (Fig. S6B). The guppy sex chromosome therefore cannot include a deleted or genetically degenerated region in which a large proportion of Y-linked genes have been lost, like that in the threespine stickleback (11). These results also exclude the possibility that presence of hemizygous sites has led to underestimated mean $F_{IS}$ values in males, and $F_{ST}$ between the sexes, in the windows analyzed. The region just proximal to 3 Mb in the female assembly, with multiple sites heterozygous for male-specific variants (Fig. 3), shows higher coverage than the flanking regions (indicate by an arrow in Fig. S6B), indicating that the male-specific SNPs are in a duplicated copy of this non-coding region.

As noted above, three individuals had genotypes differing at many LG12 sites from the fully sex-linked configuration. Two males were homozygous for the variant commonest in females, instead of being heterozygous, and one female often had the heterozygous genotype found in most males (Fig. S4). The failure of our analysis to detect regions with genotype configurations compatible with complete sex linkage is therefore not due to genotyping errors, which should affect all individuals roughly equally. Excluding these three individuals produces increased $F_{ST}$ values (Fig. S7). However, $F_{ST}$ values approaching the value expected for completely sex-linked SNPs still appear in only two of the regions showing the footprints of sex-linkage mentioned above, again not including the region of the inferred old stratum (28).

We inferred the phase of variants for LG12 sites at which no variants were present in our sample of females, and estimated a Neighbour-Joining tree (see Methods); this indicates that recombinants exist between the commonest X and Y haplotypes (Fig. S4). These haplotypes were not produced by recent recombination during captivity, as sites with genotypes incompatible with complete sex linkage are scattered throughout the entire LG12 assembly, rather than in blocks separated by evident recombination events. To investigate these individuals further, we applied principal components analysis (PCA) to our complete SNP data. Consistent with the NJ tree, Fig. S8 shows that male M2 clusters with the females on PC1, and that individuals M9 and F4 are outliers for LG12 SNPs. These two also cluster together for SNPs on LG1 and LG9 (Fig. S9), suggesting that they could be migrants from a closely related population, most likely an Aripo up-river site. Such migration has been inferred in other studies (42, 43).

*The evolutionary origin of heterochiasmy in guppies*

To test whether heterochiasmy is restricted to LG12 (and might have evolved during LG12's evolution into a sex chromosome), or is a genome-wide feature of this species, we also mapped three autosomes (1, 9 and 18). Just as for chromosome 12, crossovers were detected in most intervals tested in female meiosis (Fig. S10), but were confined to the two most distal megabases of each chromosome in male meiosis. This suggests that the guppy's heterochiasmy is genome-wide, consistent with cytogenetic experiments showing terminal localization of autosomal and sex chromosomal chiasmata in testis cells (24).

Heterochiasmy could thus have evolved in an ancestral species. We therefore attempted to infer when it arose. The closely related genus *Xiphophorus* (whose synonymous site divergence from the guppy is only about 8%), shows no major sex difference in crossover patterns (44). The same is probably also true in the medaka (*Oryzias latipes*), which can serve as a more distant outgroup for inferring changes in crossover patterns. Published male and female maps based on small numbers of sex chromosome markers suggest at most mild heterochiasmy in medaka (45).

We also used published data (46, 47) to assign crossover events in male meiosis to terminal or non-terminal regions of all medaka chromosomes (Supplementary information and Table S3). These results, together with the evidence just cited from *Xiphophorus*, suggest that the guppy's strong heterochiasmy is a derived state.

## Discussion

The guppy is excellent for studying the involvement of sexually antagonistic mutations in the evolution of sex chromosomes. Its sex chromosome pair is enriched for polymorphic male coloration genes (16) and other genes with possible functions in pigmentation processes that are potentially under SA selection (41), and also possibly genes for other sexually antagonistic traits in this highly sexually dimorphic fish. Our results can explain such an enrichment. With the genome-wide sexually dimorphic crossover pattern that we observe, a chromosome that acquires a male-determining gene will immediately be inherited as a Y chromosome that rarely recombines with the X. Closer linkage need not have evolved, and no selection against recombinant genotypes need be involved (29). This sex-linkage will facilitate subsequent establishment of SA polymorphisms since, at loci with SA polymorphisms, male-benefit alleles can maintain associations with the male-determining allele at the sex-determining locus. If recombination is very rare between an XY chromosome pair, SA polymorphisms could become established anywhere in the partially (but nearly completely) sex-linked region, even at loci physically distant from the sex-determining locus. The documented enrichment in male coloration genes does not require a concentration of genes with suitable phenotypic effects in a physically small region.

Our results contrast with the recent proposal of ongoing evolution of recombination suppression between the guppy sex chromosome pair (28), based on inferring the presence of "evolutionary strata" (see Introduction). Our detection of linkage disequilibrium with the sex-determining locus, combined with linkage analyses in our families, shows that crossing over is infrequent across the entire sex chromosome pair, except for a highly recombining pseudoautosomal region of only 1-2 Mb at the tip opposite the centromere. A model involving subsequent recombination suppression events (creating evolutionary strata) spreading the non-recombining region throughout most of the sex chromosome pair, as occurred in mammals and neognathous birds, is rejected for the guppy. The strong crossover localization in male meiosis resembles that in the frog, *Rana temporaria* (48-50), and may be more common in lower vertebrates than previously thought. Sexual dimorphism of recombination patterns, and different total genetic map lengths, have indeed been detected in several other fish species (51-55) including in an inter-species cross between sticklebacks (56), as well as in amphibians (29, 57). Heterochiasmy is known in a diversity of organisms, and several reasons for its evolution have been proposed and modelled (58). It is documented that crossover patterns change over evolutionary time (59), although closely related species appear not to have been compared.

Furthermore, our genetic mapping detected no major differences in crossover localization in sex chromosomes sampled from different natural guppy populations, including in comparisons between high- and low-predation populations. In all families studied, crossovers were strongly localized to the tip of chromosome 12. This shows that closer linkage is not currently evolving within extant populations of the species, unlike the conclusion from the previous population genomics study (28).

Crossovers in male meiosis rarely occur outside the 1-2 Mb PAR region in our mapping families, including families from both high- and low-predation natural populations (Figs.1 and S1, and Table 1). The single male recombinant individual found here yields an estimated recombination rate of 0.0022 for the entire 22 Mb centromere-proximal region; based on the Poisson distribution, the upper 95% confidence interval for the number of crossing over events in this region is one per 91 meiotic events, somewhat lower than the rate based on the four crossovers observed cytologically in this chromosome 12 region, out of 80 bivalents scored in testes of domesticated guppies (24). These observations of crossing over in males meiosis are also consistent with the rare events detected between the sex-determining locus and one partially sex-linked male colour variant (*Sb*). One recombinant was found among 1024 progeny in families of high-predation population males, and five among 387 from a low-predation site, suggesting a higher recombination rate in the latter population (17). The occurrence of X-Y crossovers, and a difference in their frequencies in the populations was further supported by observations of X-linkage of *Sb* alleles in the low-, but not the high-predation populations (17). High- and low-predation populations might therefore differ at most by the frequency and location of rare recombination events outside the PAR region. Given the rarity of such events, it will be difficult to estimate if such differences really exist.

Even very rare recombination will, however, prevent the evolution of a strongly differentiated male-specific region, and should prevent genetic degeneration, at least by Muller's ratchet, which is probably important in young sex chromosomes (60, 61). This is consistent with our genomic data, which shows no extensive male-specific region, and similar coverage of sequences in both sexes. Recombination in the centromere-proximal region should also lead to $F_{ST}$ between the sexes declining with distance from the sex-determining locus towards the centromere (34). However, we observed no such decline (Fig. 2A). This could be explained if only gene conversion, and not crossing over, occurs between the X and Y pair during male meiosis, a possibility that cannot be discounted, since gene conversion occurs in regions that do not undergo crossing over in several organisms (62). However, this cannot explain the peak in the $F_{ST}$ values observed in the region only 3 Mb from the centromere, and distant from the sex-determining region. Combined with our evidence that crossovers occur between the guppy Y and X chromosomes, this is consistent with a footprint of ongoing balancing selection on the sex chromosomes, and specifically the maintenance of SA polymorphisms at loci closely linked to the sex-determining locus; such polymorphisms lead not only to linkage disequilibrium (LD) between alleles at the locus under SA selection and the sex-determining locus, but also to LD located between closely linked neutral variants and these two loci (34). The observation of distinct Y haplotypes (Fig. S8) suggests the action of frequency-dependent selection, as such polymorphisms are unlikely to be maintained in its absence (63).

The question of how sexually dimorphic crossing over evolved is beyond the scope of the present study, as there are several possibilities, and more data are needed to distinguish between them. However, it is interesting to ask when the heterochiasmy evolved. Since localization of crossing over in male meiosis is apparently identical in all Trinidad guppy populations tested, it probably evolved before the establishment of any male coloration polymorphisms that evolved within specific guppy populations (some polymorphisms are shared by multiple populations) (17). This makes it implausible that evolutionary strata evolved independently within separate populations. Heterochiasmy is apparently genome-wide in guppies, based on our mapping of three autosomes (Fig. S10). This suggests that the pattern is not controlled by recombination modifiers located on LG12, such as inversions, but by trans-acting factors acting genome-wide. Heterochiasmy may, however, been intensified in the guppy lineage, after Y-linked male coloration polymorphisms became established. Once strong heterochiasmy evolved, more male coloration polymorphisms could have become established.

Overall, we conclude that closer linkage is not currently evolving within extant populations of this species. Rather than segregating SA polymorphisms generating ever-closer linkage with the sex-determining locus, strong heterochiasmy in guppies probably permits the maintenance of sexually antagonistic male coloration polymorphisms on the guppy Y chromosome. The

apparent absence of heterochiasmy in the related Xiphophorus species cross, and the occurrence of crossovers far from the chromosome tips in the more distantly medaka, taken together, suggest that crossover localization probably evolved recently in the guppy lineage, and it is tempting to speculate that this occurred in response to SA polymorphisms before the current sex-specific expression of these traits evolved. Future studies on more closely related fish should shed light on the time of the evolutionary change in crossover patterns, and on the involvement of sexually antagonistic mutations in the evolution of sex chromosomes.

## Methods

### Fish samples

Our study used a captive population derived from a sample collected by D.P. Croft (University of Exeter) in the lower part of the Aripo river in Trinidad (high predation zone, see Table S1 for details) in March 2008. Approximately 200 fish were used to found the population (100 males and 100 females). The fish have been maintained in a large population at Exeter, in 6 pools (3 x 1.5 x 0.6m). The pools house over 12,000 fish, and individuals have been regularly moved between pools, to maintain the absolute population size as large as possible in stock ponds. The fish used for sequencing and genetic mapping came from a colony founded at the Cornwall campus in February/March 2013 from 600 adults from this large population, and maintained with a population size of at least 1,000. To maintain a high population size, the population was split across multiple tanks, with fish regularly exchanged between tanks, and without imposing any selection or inbreeding (other than the naturally occurring selection, including through female choice). Our sample was chosen to provide sequences from 22 X chromosomes and 10 Y chromosome. Fish were humanely euthanized by terminal anaesthetic in a solution of Tricaine methanesulfonate (MS222) buffered to neutral pH with sodium bicarbonate, and preserved in EtOH at -20 C.

**Genome sequencing and analyses:** Genomic DNA was extracted from 10-12 mg of caudal tissue from 10 males and 6 females, using the DNeasy blood&tissue kit (Qiagen, Hilden, Germany) and RNase-treated according to the manufacturer's recommendations. Paired-end sequencing libraries were constructed from 3-5 µg of RNase-treated DNA and sequenced as 150-bp paired-end reads, and complete genome sequences were obtained from all 16 individuals (Table S4 provides details of the insert sizes and other statistics).

For each fastq file, reads were trimmed to remove adapters and primers, along with poor quality bases, using cutadapt (version 1.8.3), using parameters -m 35 -q 301 (64). The resulting filtered sequences were mapped to a publicly available reference female *Poecilia reticulata* genome sequence assembly (available at GenBank under accession number GCF 000633615.1). Reads were aligned against the reference genome using bwa mem, version 0.7.13 (65) with parameter -M which marks split alignments as secondary so that they can later be excluded by downstream tools. Duplicates were marked using Picard tools (version 2.8.1) (http://broadinstitute.github.io/picard) and excluded. The minimum mapping rate among the 16 guppy samples was 97.41%, and the coverage, estimated for all sites using SAMTools (v. 0.1.19), was rarely below 50 (Table S4); note that this coverage analysis differs from that of (28), who excluded male reads not perfectly matching the female assembly, in order to use lower male coverage to detect diverged Y-linked sequences). Moreover, the mapping rates were similar for males and females (Table S4).

SNPs in the resulting files in BAM format were called using the GATK pipeline (version 3.7.0). SNPs and INDELS from each sample were called using HaplotypeCaller with the following parameters: emitRefConfidence GVCF, –genotyping mode DISCOVERY, -stand call conf 30, –variant index type LINEAR, –variant index parameter 128000. Further joint genotyping was performed using GenotypeGVCFs with the default parameters, followed by use of SelectVariants (66-68) to select all variants with minimum depth of coverage (DP) of 20 and quality (QUAL) 30. VCF tools (version 0.1.15) (68) was used to generate separate VCF files for INDELS and SNPs from the initial vcf files, using parameters –keep-only-indels and –keep-indels respectively. SAMtools (v 0.1.19) was used to estimate read depths of individual sites in the assembly, for each fish, and the resulting values were analyzed using Python scripts. Coverage was high for the vast majority of sites, including those with the SNPs analyzed, in all individuals (Table S4 and Figs. S5 and S6).

The SNPs ascertained in this way were used in genetic mapping, and in the genome resequencing analyses described below.

**Genetic mapping:** A full-sib family, named LAH, of 42 individuals (25 males and 17 females) was made by crossing a single male and a single female from the Aripo population used for the genome resequencing. Genetic mapping was carried out using genotypes from several microsatellites and SNPs in genes (Table S2). For chromosome 12, the markers spanned most of the assembled chromosome, between 1 Mb and 25.6 Mb, with an average density of one marker per 1.1 Mb in this family. Microsatellites were genotyped by capillary electrophoresis using an ABI 3730 Sequencer (Applied Biosystems). SNPs from gene markers were either genotyped using RFLP assays or by direct Sanger sequencing. Genetic distances were inferred using JoinMap v. 4.0 (69), with a minimum LOD score of 3, using a regression mapping procedure, the Kosambi mapping function and a goodness-of-fit jump threshold of 5.0. In male meiosis, recombination frequencies between markers at the chromosome termini and the sets of co-segregating markers elsewhere in the physical assemblies of the chromosomes exceeded 30 cM, producing LOD scores below the threshold of 3. Linkage of these markers was supported by linkage inferred in female meiosis.

Genetic mapping was also done for seven further families, made from parents collected in natural high-predation populations in the Guanapo and Quare rivers, and low-predation populations, two from the Aripo river and one from the Quare river (Table S1). The Quare populations may represent a different species, *P. obscura* (70), but other studies have not found such populations to be differentiated (42). The family sizes are shown in Fig. 1 and Table 1.

### Population genomic analyses

Analyses were done for individual variable single nucleotide sites (SNPs) (71) that had no missing data, using VCF Tools (68). We estimated mean values for $F_{ST}$ between the two sexes $F_{IS}$ values for each sex, and Tajima's D values, in non-overlapping 50 kb windows across the physical assembly. VCF Tools was also used to estimate linkage disequilibrium (LD,) which was quantified as mean $r^2$ values between SNPs in windows, after thinning to one SNP per 2kb, and excluding variants whose minor allele frequency was below 0.2; the estimates used analyses of the diploid data from both sexes generated by sequencing, without inferring phase. For a finer scale view of heterozygote frequencies, $F_{IS}$ values were also estimated for each biallelic site, using a Python script. Principal Components Analysis (PCA) was done using the Tassel software (72).

To determine the extent of any regions exhibiting complete sex linkage, the genotype configurations of all variable sites with complete genotype information in all individuals, and no more than two alternative bases (biallelic SNPs), were examined. Such analysis allows us to identify sites with genotypes expected under complete sex linkage, i.e. sites with an allele fixed in the male population and absent in the female population, as well as sites showing partial sex linkage. Sites indicative of partial sex linkage were classified into the five categories shown in Fig. 3A. As our sample includes sequences from 22 X chromosomes, X-linkage can be readily inferred. Finally, to test whether these partially or fully sex-linked sites suggest clustering of recombinant chromosomes in certain individuals, we used a single inferred X haplotype from each female individual, and a Y haplotype from each male (Fig. S4); the tree is based on nucleotide *p*-distances analyzed in MEGA software version 7 (73).

**Data availability**: The sequence BAM files are available in the European Nucleotide Archive (STUDY_ID PRJEB22221).




1. Muller HJ (1932) Some genetic aspects of sex. *The American Naturalist* 66:118-138.
2. Muller HJ (1918) Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. *Genetics* 3:422-499.
3. Bull JJ (1983) *Evolution of Sex Determining Mechanisms* (Benjamin/Cummings, Menlo Park, CA).
4. Blackmon H & Brandvain Y (2017) Long-term fragility of Y chromosomes is dominated by short-term resolution of sexual antagonism. *Genetics* 207:1621-1629.
5. Scott M & Otto S (2017) Haploid selection favors suppressed recombination between sex chromosomes despite causing biased sex ratios. *Genetics* 207:1631-1649.
6. Rice WR (1987) The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex-chromosomes. *Evolution* 41:911-914.
7. Ponnikas S, Sigeman H, Abbot KK, & Hansson B (2018) Why Do Sex Chromosomes Stop Recombining? *Trends Genet.* 34:492-503.
8. Lahn BT & Page DC (1999) Four evolutionary strata on the human X chromosome. *Science* 286:964-967.
9. Skaletsky H, *et al.* (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825 - 837.
10. Zhou Q, *et al.* (2014) Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* 346:1332.
11. White M, Kitano J, & Peichel C (2015) Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. *Mol. Bio. Evol.* 32:1981-1995.
12. Schultheiß R, Viitaniemi H, & Leder E (2015) Spatial dynamics of evolving dosage compensation in a young sex chromosome system. *Genome Bio. Evol.* 7:581–590.
13. Bergero R, Forrest A, Kamau E, & Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* 175:1945-1954.
14. Úbeda F, Patten MM, & Wild G (2015) On the origin of sex chromosomes from meiotic drive. *Proceedings of the Royal Society. Series B: Biological Sciences* 282:20141932.
15. Matsumoto T, Yoshida K, & Kitano J (2017) Contribution of gene flow to the evolution of recombination suppression in sex chromosomes. *J. Theoret. Biol.*
16. Lindholm A & Breden F (2002) Sex chromosomes and sexual selection in Poeciliid fishes. *Am. Nat.* 160:S214-S224.
17. Haskins C, Haskins EF, McLaughlin J, & Hewitt RE (1961) Polymorphisms and population structure in *Lebistes reticulatus*, an ecological study. *Vertebrate Speciation*, ed Blair WF

(University of Texas Press, Austin, TX), pp 320-395.
18. Gordon SP, López-Sepulcre A, Rumbo D, & Reznick DN (2017) Rapid changes in the sex linkage of male coloration in introduced guppy populations. *Am. Nat.*
19. Endler JA (1980) Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34:76-91.
20. Magurran AE (2005) *Evolutionary Ecology: The Trinidadian Guppy* (Oxford University Press, Oxford).
21. Olendorf R, *et al.* (2006) Frequency-dependent survival in natural guppy populations. *Nature* 441:633-636.
22. Künstner A, *et al.* (2017) The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. *PLOS ONE* 11:e0169087.
23. Nanda I, *et al.* (2014) Sex chromosome polymorphism in guppies. *Chromosoma* 123:373-383.
24. Lisachov A, Zadesenets K, Rubtsov N, & Borodin P (2015) Sex chromosome synapsis and recombination in male guppies. *Zebrafish* 12:174-180.
25. Hughes K, Houde A, Price A, & Rodd F (2013) Mating advantage for rare males in wild guppy populations. *Nature* 503:108–110.
26. Connallon T & Knowles LL (2005) Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet.* 21:495-499.
27. Gordon SP, López-Sepulcre A, & Reznick DN (2012) Predation-associated differences in sex-linkage of wild guppy coloration. *Evolution* 66:912-918.
28. Wright A, *et al.* (2017) Convergent recombination suppression suggests a role of sexual conflict in guppy sex chromosome formation *Nature Communications* 8:14251.
29. Rodrigues N, Studer T, Dufresne C, & Perrin N (2018) Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Mol. Bio. Evol.* 35:942-948.
30. Jeffries DL, *et al.* (2018) A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nature Communications* 9:4088.
31. Tripathi N, Hoffmann M, Weigel D, & Dreyer C (2009) Linkage analysis reveals the independent origin of Poeciliid sex chromosomes and a case of atypical sex inheritance in the guppy (*Poecilia reticulata*). *Genetics* 182:365–374.
32. Tripathi N, *et al.* (2009) Genetic linkage map of the guppy, *Poecilia reticulata*, and quantitative trait loci analysis of male size and colour variation. *Proceedings of the Royal Society B* 276:2195-2208.
33. Shen X, *et al.* (2007) Construction of genetic linkage maps of guppy (*Poecilia reticulata*) based on AFLP and microsatellite DNA markers. *AQUACULTURE* 271:178-187.
34. Kirkpatrick M & Guerrero R (2014) Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197:531-541.
35. Darolti J, Wright AE, Pucholt P, Berlin S, & Mank JE (2018) Slow evolution of sex-biased genes in the reproductive tissue of the dioecious plant *Salix viminalis Mol. Ecol.* 27:694–708.
36. Cruickshank T & Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23:3133-3157.
37. Wilson C, *et al.* (2014) Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* 198:291-1308.
38. Davidson W, Huang T, Fujiki K, Schalburg Kv, & Koop B (2009) The sex determining loci and sex chromosomes in the family Salmonidae. *Sexual Development* 3:78-87.
39. Kamiya T, *et al.* (2012) A trans-species missense SNP in *Amhr2* is associated with sex determination in the Tiger Pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genetics* 8:e1002798.
40. Bao L, *et al.* (2019) The Y chromosome sequence of the channel catfish suggests novel sex determination mechanisms in teleost fish. *BMC Biology* 17.
41. Morris J, Darolti I, Bloch N, Wright A, & Mank J (2018) Shared and species-specific patterns of nascent Y chromosome evolution in two guppy species. *Genes* 9:238.
42. Willing E-M, *et al.* (2010) Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Mol. Ecol.* 19:968–984.
43. Barson N, *et al.* (2009) Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source-sink metapopulation structure, founder events and population bottlenecks. *J. Evol. Biol.* 22:485-497.
44. Walter RB, *et al.* (2004) A microsatellite genetic linkage map for *Xiphophorus*. *Genetics* 168:363-372.
45. Kondo M, Nagao E, Mitani H, & Shima A (2001) Differences in recombination frequencies during male and female meioses of the sex chromosomes of the medaka, *Orzyias latipes*. *Genet. Res.* 78:23-30.
46. Nishimura T, *et al.* (2014) Analysis of a novel gene, *Sdgc*, reveals sex chromosome-dependent differences of medaka germ cells prior to gonad formation. *Development* 141(17):3363-3369.
47. Kimura T, Shinya M, & Naruse K (2012) Genetic analysis of vertebral regionalization and number in medaka (*Oryzias latipes*) inbred lines. *G3-Genes Genomes Genetics* 2:1317-1323.
48. Brelsford A, Dufresnes C, & Perrin N (2015) High-density sex-specific linkage maps of a European tree frog (*Hyla arborea*) identify the sex chromosome without information on offspring sex. *Heredity* 116:177-181.
49. Brelsford A, Rodrigues N, & Perrin N (2016) High-density linkage maps fail to detect any genetic component to sex determination in a *Rana temporaria* family. *J. Evol. Biol.* 29:220-225.
50. Rodrigues N, *et al.* (2017) *Dmrt1* polymorphism and sex chromosome differentiation in *Rana temporaria*. *Mol. Ecol.* 26(19):4897-4905.
51. Singer A, *et al.* (2002) Sex-specific recombination rates in zebrafish (*Danio rerio*). *Genetics* 160:649–657.
52. Kai W, *et al.* (2011) Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in teleosts and mammals. *Genome Bio. Evol.* 3:424-442.
53. Kai W, *et al.* (2005) A genetic linkage map for the Tiger Pufferfish, *Takifugu rubripes*. *Genetics* 171:227-238.
54. Sakamoto T, *et al.* (2000) A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155:1331-1345.
55. Lien S, *et al.* (2011) A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics* 12:615.
56. Sardell J, *et al.* (2018) Sex differences in recombination in sticklebacks. *G3-Genes Genomes Genetics* 8:1971-1983.
57. Berset-Brändli L, Jaquiéry J, Broquet T, Ulrich Y, & Perrin N (2008) Extreme heterochiasmy and nascent sex chromosomes in European tree frogs. *Proceedings of the Royal Society of London B.* 275:1577-1585.
58. Lenormand T (2003) The evolution of sex dimorphism in recombination. *Genetics* 163:811-822.
59. Dutheil J, Munch K, Mailund T, & Schierup M (2015) Strong selective sweeps on the x chromosome in the human-chimpanzee ancestor explain its low divergence. *PLoS Genetics* 11:e1005451.
60. Charlesworth D, Morgan MT, & Charlesworth B (1993) Mutation accumulation in finite outbreeding and inbreeding populations. *Genet. Res.* 61:39-56.
61. Blaser O, Grossen C, Neuenschwander S, & Perrin N (2013) Sex-chromosome turnovers induced by deleterious mutation load. *Evolution* 67:635-645.
62. Korunes K & Noor M (2017) Gene conversion and linkage: effects on genome evolution and speciation. *Mol. Ecol.* 26:351-364.
63. Clark A (1987) Natural selection and Y-linked polymorphism. *Genetics* 115:569-577.
64. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10-12.
65. Li H & Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595.
66. DePristo M, *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43:491–498.
67. Auwera G, *et al.* (2013) From fastq data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. 10-11, 10.11-33. *Current Protocols in Bioinformatics* 43(10-11):1-33.
68. Danecek P, *et al.* (2011) The variant call format and vcftools. *Bioinformatics* 27:2156–2158.
69. van Ooijen JW (2006) *JoinMap 4.0 Software for the Calculation of Genetic Linkage Maps in Experimental Populations* (Plant Research International, Wageningen, the Netherlands:).
70. Schories S, Nanda I, & Schartl M (2009) Description of *Poecilia* (Acanthophacelus) *obscura* n.sp., (Teleostei: Poeciliidae), a new guppy species from western Trinidad, with remarks on *P.wingei* and the status of the 'Endler's guppy'. *Zootaxa* (2266):35–50.
71. Hudson RR, Boos DD, & Kaplan NL (1992) A statistical test for detecting geographic subdivision. *Mol. Bio. Evol.* 9:138-151.
72. Glaubitz J, *et al.* (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*:e0090346.
73. Kumar S, Stecher G, & Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 for bigger datasets. *Mol. Bio. Evol.* 33:1870–1874.
74. Hill WG & Robertson A (1968) Linkage disequilibrium in finite populations. *Theoret. Appl. Genet.* 38:226-231.

**Acknowledgements** : We thank two anonymous reviewers for their help in improving this manuscript, the University of Exeter animal technicians for their support with fish husbandry, and colleagues Konrad Lohse, Darren Obbard, and Alex Twyford at Edinburgh, and Alastair Wilson at Exeter, for discussions and comments on the manuscript. The project was supported by ERC grant number 695225 (GUPPYSEX). Genome resequencing and bioinformatics analyses (sequence alignments and production of BAM and VCF files) were carried out by Edinburgh Genomics, The University of Edinburgh. Edinburgh Genomics is partly supported through core grants from NERC (R8/H10/56), MRC (MR/K001744/1) and BBSRC (BB/J004243/1).

**Contributions** DC designed the study with input from RB. LY provided the samples guppies with known sexes, and produced the families used for genetic mapping. JG and BB performed all molecular biology experiments, including genotyping individuals for genetic mapping. RB analyzed the data with input from DC. RB and DC supervised all aspects of the study. RB and DC collaborated to write the paper. **Corresponding authors**: Correspondence to Roberta Bergero or Deborah Charlesworth

**Contributions**
DC designed the study with input from RB. LY provided the samples guppies with known sexes, and produced the families used for genetic mapping. JG and BB performed all molecular biology experiments, including genotyping individuals for genetic mapping. RB analyzed the data with input from DC. RB and DC supervised all aspects of the study. RB and DC collaborated to write the paper.
**Corresponding authors**: Correspondence to Roberta Bergero or Deborah Charlesworth